#2

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office

September 30, 2003

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: ***60/406,757***
FILING DATE: ***August 28, 2002***
RELATED PCT APPLICATION NUMBER: *PCT/US03/26827*

CD DISK IS THE APPLICATION FOR THE ABOVE REFERENCED INFORMATION

By Authority of the
COMMISSIONER OF PATENTS AND TRADEMARKS

N. WILLIAMS
Certifying Officer

Due to the format in which this priority document and accompanying certification was transmitted, the International Bureau has been unable to upload these documents in full into its information systems. Only a copy of the certification page and the first four pages of the specification are enclosed herewith. The International Bureau will provide a complete copy f this document upon request.

60406757 .092802

PATENT APPLICATION
ATTY REF NO. 19195.003

CONSERVED AND SPECIFIC STREPTOCOCCAL GENOMES

FIELD OF THE INVENTION

The invention relates to polynucleotides which are conserved or specific to one or more species of Streptococcus, Streptococcus species serotypes, and/or serotype isolates. The conserved or specific genomic regions can be used to identify, screen and develop vaccines and other treatments for Streptococcal infections and can be used in diagnostic assays to diagnose and identify Streptococcal infections.

BACKGROUND OF THE INVENTION

The genus *Streptococcus* consists of Gram-positive, chain-forming, spherical bacterial cells. Three species of clinical interest are *S.pneumoniae* ("pneumococcus" or "S.pn."), *S.pyogenes* ('group A streptococcus' or 'GAS') and *S.agalactiae* ('group B streptococcus' or 'GBS'). Infections with these three pathogenic streptococci lead to conditions including pharyngitis, toxic shock syndrome and necrotizing fasciitis.

Once thought to infect only cows, GBS is now known to cause serious disease, bacteraemia and meningitis in immunocompromised individuals and neonates. There are two known types of neonatal infection. The first (early onset, usually within 5 days of birth) is manifested by bacteraemia and infection. It is generally contracted vertically as a baby passes through the birth canal. GBS is thought to colonize the vagina of about 25% of young women; approximately 1% of infants born via a vaginal birth to colonised mothers will become infected. Mortality resulting from these infections is between 50 – 70%. The second type of neonatal infection is a meningitis that occurs 10 to 60 days after birth. If pregnant women are vaccinated with type III capsule so that the infants are passively immunised, the incidence of the late onset meningitis is generally reduced, although not entirely eliminated.

The "B" in "GBS" refers to the Lancefield classification, which is based on the antigenicity of a carbohydrate which is soluble in dilute acid and called the C carbohydrate. Lancefield identified 13 types of C carbohydrate, designated A to O, that could be serologically differentiated. The organisms that most commonly infect humans are found in groups A, B, D,

and G. Within group B, strains can be divided into at least 9 serotypes (Ia, Ib, II, III, IV, V, VI, VII, and VIII) based on the structure of their polysaccharide capsule. Further categories based on, for example, the expression of certain proteins have also been developed.

GBS strains of polysaccharide capsule Type V were rarely isolated before the mid-1980's but now account for approximately one-third of clinical isolates in the US. Type V is the most common capsular serotype associated with invasive infection in nonpregnant adults, and the emergence of Type V strain over the past decade has been temporarily linked to an increase in GBS disease in this population.

Group A streptococcus is a frequent human pathogen, estimated to be present in between 5 – 15% of normal individuals without signs of disease. When host defences are compromised, or when the organism is able to exert its virulence, or when it is introduced into vulnerable tissues or hosts, however, an acute infection occurs. Diseases include puerperal fever, scarlet fever, erysipelas, pharyngitis, impetigo, necrotising fasciitis, myositis and streptococcal toxic shock syndrome.

Pneumococcus is the most common cause of acute respiratory infection and otitis media and is estimated to result in over 3 million deaths in children every year worldwide from pneumonia, bacteremia, or meningitis. Even more deaths occur among elderly people, among whom S. pn. is the leading cause of community-acquired pneumonia and meningitis. Since 1990, the number of penicillin-resistant strains has increased from 1 to 5% to 25 to 80% of isolates, and many strains are now resistant to commonly prescribed antibiotics such as penicillin, macrolides, and fluoroquinolones. *See* Tettelin, et al. (2001) *Science* 293, 248-506.

The complete genomic sequence of a virulent isolate of *S. pneumoniae* was published by Tettelin, et al. (2001) *Science* 293, 248-506 and is available at the TIGR website at http://www.tigr.org. as well as on GEN BANK (available through the Pub Med website at http://www.ncbi.nlm.nih.gov/entrez/query.fcgi). The genomic sequence, the Tettelin article and its published supplemental material are incorporated herein by reference in their entirety.

The complete genomic sequence of an M1 strain of *S. pyrogenes* was published by Ferretti, et al. (2001) *Proc. Natl. Acad. Sci. USA* 98, 4658 – 4663 and is available at the TIGR

website at http://www.tigr.org. The genomic sequence, the Ferretti article and its published supplemental materials are incorporated herein by reference in their entirety.

The complete genomic sequence of a serotype V strain of S. agalactiae (type V strain 2603 V/R) is published on the date of this filing, August 28, 2002 by Gen Bank Accession no. AE009948 (available through Pub Med at http://www.ncbi.nlm.nih.gov/entrez/query.fcgi and/or is available on the same day at the TIGR website at http://www.tigr.org. This sequence has been downloaded from the TIGR website and is included in this application as Table 39. Most of this sequence is also availabe in PCT International Patent Application Publication WO 02/34771. The genomic sequence, the Tettelin article and its published supplemental materials are incorporated herein by reference in their entirety.

Current treatments for *Streptococcal* infections include both antibiotics and prophylactic vaccination. Current vaccines, particularly with respect to GBS, suffer from poor immunogenicity, while the emergence of antibiotic resistant strains has lessened the effectiveness of currently used antibiotics. Accordingly, there is an increasing need for the development of new vaccines and antibiotics (as well as other small molecule bacterial inhibitors) to help prevent and treat Streptococcal infections.

Applicants have identified regions of the Streptococcal genomes which can be used to identify and develop new vaccines and treatments for Streptococcal infections. Specifically, Applicants have identified polynucleotides of the Streptococcal genome which are conserved or specific to Streptococcal species, species serotypes, and/or specific serotype isolates. These polynucleotides and their expressed polypeptides can be used to screen, develop and design new vaccines, antibiotics and other small molecule bacterial inhibitors. These polynucleotides and their expressed polypeptides can further be used to diagnose and identify Steptococcal infections.

SUMMARY OF THE INVENTION

The invention relates to polynucleotides which are conserved or specific to one or more species of Streptococcus, Streptococcus species serotypes, and/or serotype isolates. In particular, the invention relates to polynucleotides from Streptococcus which are conserved or specific to one or more of the species of *S. pneumoniae* ("pneumococcus" or "S. pn."), *S. pyogenes* ("group

A streptococcus" or "GAS"), and *S. agalactiae* ("group B streptococcus" or "GBS"). The invention further relates to polynucleotides which are conserved or specific to one or more Streptococcal species serotypes, such as GBS serotypes Ia, Ib, II, III, IV, V, VI, VII, and VIII. The invention still further relates to polynucleotides which are conserved or specific to one or more clinical isolates of a Streptococcus species.

The invention is based on the identification of the following Subsets of genes. Genes falling within each subset are described with respect to referenced tables, lists, and/or figures (in particular the CGH map depicted in Figure 1).

The following Subsets related to the GBS genome:

GBS Subset 1: 1060 GBS genes which have homologs with GAS and with pneumococcus (Table 8);

GBS Subset 2: 225 GBS genes which have homologues with GAS, but not with pneumococcus (Table 10);

GBS Subset 3: 176 GBS genes which have homologues with pneumococcus but not with GAS (Table 9);

GBS Subset 4: 683 GBS genes which do not have homologues with GAS or pneumococcus (specific to GBS vs GAS and pneumococcus) (Table 11).

The invention is based on the identification of the following subsets of genes within the GAS genome:

GAS Subset 1: 1006 GAS genes which have homologues with GBS and with pneumococcus (Table 33);

GAS Subset 2: 212 GAS genes which have homologues with GBS but do not have homologues with pneumococcus (Table 34);

GAS Subset 3: 62 GAS genes which have homologues with pneumococcus but do not have homologues with GBS (Table 35);

GAS Subset 4: 416 GAS genes which do not have homologues with either GBS or pneumococcus. This Subset can be determined by subtracting the above subsets from the published genome.